Has DRG Creep Crept Up?

Decomposing the Case Mix Index Change Between 1987 and 1988

Grace M. Carter, Joseph P. Newhouse,
Daniel A. Relles

RAND

The research described in this report was supported by the Health Care Financing Administration, U.S. Department of Health and Human Services, Cooperative Agreement 99-C-98489/9-06, and by the Prospective Payment Assessment Commission.

ISBN: 0-8330-1195-2

Published 1991 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

R-4098-HCFA/ProPAC

Has DRG Creep Crept Up?

Decomposing the Case Mix Index Change Between 1987 and 1988

Grace M. Carter, Joseph P. Newhouse,
Daniel A. Relles

Supported by the
Health Care Financing Administration,
U.S. Department of Health and Human Services
Prospective Payment Assessment Commission

RAND

RAND/UCLA/Harvard
Center for Health Care
Financing Policy Research

PREFACE

This report describes the results of a project conducted to determine how much of the increase in the Medicare Prospective Payment System's case mix index from 1987 to 1988 was caused by the resource needs of patients rather than by changes in coding and administrative practices. The project was performed within the RAND/UCLA/Harvard Center for Health Care Financing Research under a cooperative agreement between the Center and the Health Care Financing Administration (HCFA). A portion of the supporting funds was provided to the HCFA by the Prospective Payment Assessment Commission. The findings should be of interest to policymakers, analysts, and hospital administrators concerned with the Medicare rate-setting process.

SUMMARY

Between 1987 and 1988, the Medicare program's case mix index (CMI), which is an average of the diagnosis-related group (DRG) weights of hospital patients, increased by about 3.3 percent. Since the Health Care Financing Administration's (HCFA's) payments to hospitals are roughly proportional to the CMI, the increase in the CMI meant that the HCFA paid hospitals approximately $1.5 billion more than it would have if the CMI had not increased.

We undertook a project to determine how much of the 1987–1988 CMI increase was caused by an increase in the real resource needs of Medicare patients and how much was caused by changes in coding and administrative practices.

The beginning of 1988 saw the largest change yet in the definition of DRGs and in the Grouper program that assigns cases to DRGs based on hospital-coded values for diagnoses and procedures. It became important for hospitals to code complicating diagnoses and comorbidities for patients who were over 69 because, in many cases, this coding would cause assignment to a higher-weighted DRG. Also, new DRGs with very high weights were established for patients who received a tracheostomy or who were on ventilator support, providing an incentive for hospitals to more completely code these services.

METHOD

The essence of our method was to have coders at the SuperPRO, the entity responsible for reviewing the performance of Medicare's Professional Review Organizations (PROs), recode a random sample of charts from 1987 at approximately the same time they were coding a random sample of charts from 1988. The difference between the CMIs produced by these codings is a measure of true change in resource needs, since each year's patient discharges are being subjected to the same coding process. This coding also provides a consistent standard against which to measure changes in hospital coding behavior.

Since coding and administrative change can itself be decomposed into various types of effects, we decomposed the CMI increase into the following four components, the first three of which were measured using the 1988 Grouper program:

1. The increase attributable to true change, measured as the CMI change based on SuperPRO coding using the same coding standards for each year's discharges. The HCFA would like the CMI to increase only by this much.
2. The increase attributable to coding changes common to the SuperPRO and hospital, measured as the CMI change for 1987 cases between the SuperPRO's latest coding (i.e., using 1988 standards) and its original coding of the same cases.
3. The increase attributable to hospital-specific coding changes, measured as the difference between the SuperPRO and hospital codings of a given case using the standards of the year of discharge.
4. The increase due to the Grouper normalization effect, measured as the difference between the CMI measured by the 1988 Grouper for 1987 discharges and the CMI measured by the 1987 Grouper for the same discharges. The year in which the normalization effect is measured is chosen so that the sum of all four components of change equals total change.

The complete decomposition was estimated for a sample of 23 states for which SuperPRO data were available for fiscal year (FY) 1988. These states covered all regions of the United States and had approximately 56 percent of the country's Medicare discharges. Data on all states were available for FY 1987. We compared the coding behavior for the 1987 discharges in the 23 sample states to that for the nation as a whole and found them to be quite similar. Under the assumption that hospital coding behavior in the 23 sample states was also similar to national coding behavior in FY 1988, we decomposed the national CMI increase as well.

The SuperPRO data base does not contain outlier cases. To adjust our estimates to account for the missing outliers, we built regression models that estimate SuperPRO coding as a function of DRG characteristics. We then used the characteristics of outlier cases found on the Medicare Provider Analysis and Review (MEDPAR) file to estimate the coding that the SuperPRO would have used for outlier cases. A generalized least squares (GLS) methodology was used to reconcile our sample estimates with the amount of change estimated for the population from the MEDPAR file.

RESULTS AND CONCLUSIONS

In both FY 1987 and FY 1988, the DRG assigned by the SuperPRO agreed with the DRG assigned by the hospital in roughly 85 percent

of the cases. Somewhat surprisingly, the SuperPRO assigned a lower DRG weight than the hospital in only 0.5 percent more cases than it assigned a higher DRG weight. Because of this high rate of agreement between the hospital and SuperPRO assignments, the 1987 Grouper used on FY 1987 patient discharges gave the SuperPRO a CMI similar to the CMI that was paid for the sample cases. (The CMIs were within 0.1 percent of each other.)

The results of our comparison of hospital and SuperPRO coding using the 1988 Grouper were quite different. For FY 1987 discharges, the SuperPRO's original coding resulted in 1988 Grouper DRGs with significantly higher weights than those of the coding of either the hospitals or PROs. Most of this difference had disappeared by 1988, when the 1988 Grouper was in use. Hospital coding changes that were not duplicated at the SuperPRO accounted for about 20 percent of the CMI increase.

There were also measurable differences between the SuperPRO recoding of 1987 cases using the FY 1988 coding rules and its original coding of those cases. The change in coding practice that occurred at the SuperPRO almost certainly occurred at hospitals also, since the 1988 coding of the two agencies was so similar. This coding change accounted for about 30 percent of the national CMI increase.

The coding changes that occurred at both the SuperPRO and the hospitals were directly related to changes in the Grouper program introduced in FY 1988. Two kinds of cases were responsible for 93 percent of the increase in the CMI from the first coding to the recoding. The first kind, ventilator support cases, was responsible for over 33 percent of the increase. In the first coding, which was performed to verify the accuracy of PRO DRG assignment under the 1987 Grouper, ventilator support had no effect on DRG accuracy and was rarely coded. In the recoding, however, ventilator support was noted much more frequently because it changed DRG assignment under the 1988 Grouper. The MEDPAR data recorded a similar, tenfold increase from 1987 to 1988 in the frequency of ventilator support coding.

The second kind of cases that led to the increase in CMI from the first coding to the recoding was the *paired DRGs*, with the members sharing a primary diagnosis and procedures but separated by the presence of complications or comorbidities for one DRG. Paired DRGs accounted for 56 percent of the 3600 cases that were recoded nationally. Excluding ventilator support cases, the CMI increase from these paired DRGs averaged six times the increase from nonpaired, nonventilator cases.

The most important part of our CMI increase decomposition is the differentiation of true change from change caused by the sum of the coding and administrative components. For the sample states, our point estimate is that true change was responsible for 50 percent of the total CMI increase. Although the point estimates are imprecise because of the small size of the sample, we can be sure that coding changes and the Grouper normalization change contributed to the CMI's rate of increase between 1987 and 1988. Our analysis of the sample data showed that hospitals changed their coding behavior relative to a consistent coding standard.

Our national decomposition resulted in a similar division of the CMI increase into true change and change caused by coding/administrative factors. Our point estimate of the annual rate of increase in the national CMI caused by factors other than the real needs of patients is 1.7 percent, which translates into an additional annual expenditure of approximately $800 million in FY 1988 and FY 1989.

The HCFA could develop a method for ongoing monitoring of national increases in the CMI. The logical structure of our research shows the important elements of the analysis. All that is needed to implement CMI monitoring is a data base containing a nationally representative set of cases. As currently structured, the SuperPRO data base is not adequate for this purpose. Consequently, if the HCFA wishes to have an ongoing annual procedure for monitoring the sources of CMI change, it will have to collect the necessary data now. Because the amount of DRG creep is an important factor in determining the rate of increase in the PPS payment rate, it seems to us that the HCFA should have this capability.

ACKNOWLEDGMENTS

This research would not have been possible without the efforts of the SuperPRO staff at SysteMetrics, Inc., in Santa Barbara, California. We would like to thank their medical records technicians, physicians, data coders, and programmers. We also thank Cathleen Barnes, who ably organized and supervised the work of the SysteMetrics staff.

We thank Timothy Greene of the Health Care Financing Administration (HCFA) and Stuart Guterman of the Prospective Payment Assessment Commission (ProPAC) for providing very useful comments on an earlier version of this report. The report also benefited from very helpful reviews by RAND staff members Susan Marquis and Emmett Keeler. We thank Jeri O'Donnell for editing the manuscript, and we are grateful to the HCFA and to the ProPAC for supporting this work.

CONTENTS

TABLES

1. INTRODUCTION

The Medicare program pays each hospital approximately in proportion to the hospital's case mix index (CMI), a simple average of the diagnosis-related group (DRG) weights of its patients. Since the introduction of Medicare's Prospective Payment System (PPS), the national CMI has increased each year, as shown in Table 1.1. Because each 1 percent increase in the national CMI represents about an additional $450 million of federal expenditure, CMI changes have been controversial.[1]

Hospitals have argued that CMI increases largely reflect technical changes that lead physicians to attempt more complex, highly weighted treatments (e.g., more cardiovascular surgery) for the elderly, as well as the shift of low-weighted cases (e.g., cataract operations) out of the hospital into outpatient treatment facilities. The government, however, has been concerned that CMI changes may be a reflection solely of changes in coding practices—i.e., similar sets of patients and treatments being coded into higher-weighted categories, a phenomenon some people dub *DRG creep* (Simborg, 1981). The Medicare program should ideally adjust its payments in response to the part of the CMI change that is *true* change (i.e., reflective of the real resource needs of patients) and not in response to the part that is merely the result of coding changes.

Table 1.1

Change in CMI by Year

Year	CMI Increase (%)
1984–1985	4.2
1985–1986	3.0
1986–1987	2.4
1987–1988	3.3

NOTE: Data restricted to PPS cases.

[1]The Prospective Payment Assessment Commission (ProPAC) estimated that PPS payments would equal $45 billion in fiscal year (FY) 1989 (ProPAC, 1989, Table 5.1).

BACKGROUND

Conditional upon a given coding system, one would expect that changes in the CMI would converge over time to the true rate of change in the CMI. Hospitals, after all, should learn over time how to code their patients to maximum advantage, and once all hospitals are maximally upcoding (but not fraudulently coding), further changes in the CMI would represent true change.

As shown in Table 1.1, what happened in the first three years of the PPS program conformed to the convergence hypothesis: the CMI rate of increase fell each year. The 1987–1988 change, however, appeared to conflict with the hypothesis: the CMI rate of increase jumped by more than a third. The convergence hypothesis is conditional upon a given coding system, however, and the beginning of FY 1988 saw the largest change yet in the Medicare coding system. The Grouper program, which maps diagnoses and procedures into DRGs, underwent many changes in 1988, two of which turned out to be significant:[2]

1. Whereas patients with certain diagnoses and procedures had previously been reimbursed at a higher rate if they (a) also had a relevant complicating condition or comorbidity or (b) were over 69, beginning in 1988, only those with the complications/comorbidities were to be reimbursed at the higher rate. Moreover, the spread in weight increased for the two members of each of such *DRG pair* (e.g., Peripheral Vascular Disorders with Complication or Comorbidity and Peripheral Vascular Disorders with No Complication or Comorbidity).
2. New DRGs were established for patients who received a tracheostomy or who were on ventilator support. Although few patients were in these new DRGs in FY 1988 (0.1 and 0.4 percent, respectively), the proportion of patients with these characteristics was greater than it had been in previous years, and these DRGs were accorded very large weights (11.8772 and 3.1757).

It is therefore reasonable to believe that much of the jump in the CMI's rate of increase between 1987 and 1988 might be attributable to coding effects. However, one cannot rule out the possibility that the rate of true change also rose between the two years.

---

[2]Other changes included revisions of the underlying coding system for several diagnoses and procedures, the reassignment of primary diagnoses to DRGs in several major diagnostic categories (MDCs), refinement of the list of complications and comorbidities, and revision of the surgical hierarchy. (See Department of Health and Human Services, 1987.)

The total CMI change in 1988 represented roughly $1.5 billion in additional payments by the Medicare program that year. This sum of money put in sharp relief the question, how much of these payments represented true CMI change and how much represented DRG creep? We therefore undertook a study to decompose the change in the CMI into true change and change caused by coding/administrative factors.

A year ago we undertook the task of decomposing the 1986–1987 CMI change into the portion caused by true change and the portion caused by DRG creep (Carter, Newhouse, and Relles, 1990). That study used the random sample of records gathered by the *SuperPRO*, the private firm that independently recodes a sample of the charts to which hospitals and professional review organizations (PROs) have assigned DRGs. The Health Care Financing Administration (HCFA) contracts with a PRO in each state to review the quality and appropriateness of care given to Medicare beneficiaries. As part of its duties, the PRO recodes a random sample of each hospital's cases to verify the accuracy of the hospital's coding. The intent of the SuperPRO's recoding is to check the PRO's accuracy in reviewing hospital coding. An expert coder at the SuperPRO reads the medical chart independently of the hospital and PRO to determine the correctness of the DRG assigned by the hospital and PRO. By having the expert coder assign DRGs to random charts drawn from different time periods, it is possible to arrive at a standard for true CMI change.

We found that about two-thirds of the 2.4 percent change in CMI from 1986 to 1987 was true; that is, the true rate of increase in required hospital resources caused by changes in the distribution of cases across DRGs was around 1.6 percent.

The methods we used for the study being reported here are similar to those of the earlier study. But, as we describe in detail below, our new results are more uncertain, partly because our sample was smaller. Both studies relied on the random sample of charts available at the SuperPRO. Through 1987 these charts were available for every state. In 1988, however, they were available for only 23 states and were fewer in number.

The current study also differs from the previous one in another way. The details of our decomposition are conditional upon a given Grouper program. In decomposing the 1986–1987 change, we found that the results were similar whether we used the 1986 or the 1987 Grouper program. Unfortunately, but not surprisingly given the change in the Grouper program in 1988, our decomposition of coding and administrative effects for the 1987–1988 change differed for the 1987 and 1988 Groupers, although the sums of all coding and admin-

istrative effects for the two Groupers were reasonably similar. In the body of the text, we present only the decomposition measured by the FY 1988 Grouper. Hospitals may have changed their coding behavior in ways that can be measured only with the 1988 Grouper. In addition, the FY 1988 Grouper was a substantial refinement of the DRG system; it increased the correlation between DRG weight and resource use and thus should be a better instrument for measuring true CMI change. A decomposition using the 1987 Grouper (presented in App. A) shows roughly the same amount of real change.

DECOMPOSITION

The essence of our method is to have coders recode a random sample of charts from an earlier year (in this case 1987) at approximately the same time they are coding a random sample of charts from the current year (1988). The difference between these measures is a measure of true change in resource needs, since each year's discharges are being subjected to the same coding process. This coding also provides a consistent standard against which to measure changes in hospital coding behavior.

We now show how we formally decomposed the measured increase in the CMI. It is important for the reader to master this decomposition, because our later results are keyed to it. Certain components in this decomposition were measured with the 1988 Grouper. Appendix A describes how the decomposition changed when we measured these components with the 1987 Grouper.

The symbols used here are as follows: CMI88 and CMI87 are the CMIs for discharges in 1988 and 1987, respectively; H represents the combined coding of the hospital and PRO; S represents the SuperPRO coding; C represents current-year coding (either 1987 or 1988); and R represents the recoded records.

1. Measured change =

$$\text{CMI88(H,C,G88)} - \text{CMI87(H,C,G87)},$$

   or the change in the CMI from 1987 to 1988 as coded for payment using the current year's standards and the current year's Grouper to assign DRGs and weights. This is the numerator for the percentages given in Table 1.1.

2. True change =

$$\text{CMI88(S,R,G88)} - \text{CMI87(S,R,G88)},$$

or the change in the CMI from 1987 to 1988 as recoded by the SuperPRO using the same coding standards and the 1988 Grouper. The HCFA would like payments to increase only this much (in the absence of policy decisions).

3. Coding change common to SuperPRO and hospital =

$$\mathrm{CMI87(S,R,G88)} - \mathrm{CMI87(S,C,G88)},$$

or the change in the CMI for 1987 cases when the SuperPRO recoded using later standards.

4. Hospital-specific coding change =

$$[\mathrm{CMI88(H,C,G88)} - \mathrm{CMI88(S,C,G88)}] - [\mathrm{CMI87(H,C,G88)} - \mathrm{CMI87(S,C,G88)}],$$

or the change in the difference between hospital coding and SuperPRO coding. This is the change resulting from changes in hospital coding standards that did not occur at the SuperPRO.

5. Grouper normalization effect =

$$\mathrm{CMI87(H,C,G88)} - \mathrm{CMI87(H,C,G87)}.$$

This effect is due to the difference in the CMI measured by the 1987 and 1988 Groupers in 1987, at the start of the period. The 1988 Grouper was calibrated so that it would produce the same CMI on FY 1986 cases as the 1987 Grouper did.[3] By FY 1987, the CMIs produced by the two Groupers diverged. From the viewpoint of the 1987–1988 CMI increase, the Grouper normalization effect is purely an administrative phenomenon due to misnorming of the 1988 Grouper.

6. By an identity,

$$\mathrm{CMI88(S,C,G88)} - \mathrm{CMI88(S,R,G88)} = 0,$$

because the SuperPRO coded the 1988 cases contemporaneously (approximately); that is, no recoding was done for 1988.

---

[3]The 1986 data were the latest figures available at the deadline for publishing the 1988 weights.

If we add together Eqs. 2 through 6, we obtain the expression for measured change in Eq. 1. The principal task of our research was to decompose the change measured in Eq. 1 into the components described in Eqs. 2 through 5, especially the component of Eq. 2 versus the components of Eqs. 3 through 5.

We used a sample of SuperPRO cases from each year to yield an unbiased estimate of the total CMI change in the 23 sample states and the decomposition of the total CMI change in those states. Because of the small size of the sample, however, we obtained a much more precise estimate of the total CMI change by using data from the Medicare Provider Analysis and Review (MEDPAR) file and adjusted our decomposition, as explained below.

2. DATA AND METHODS

This section begins with a description of the SuperPRO sample that we used to measure true change in the CMI and changes resulting from coding and administrative factors. Data from the MEDPAR file, which the HCFA constructs for analyses of Medicare inpatient stays and which we used to measure total change, are described next. Approximately 5 percent of cases were designated as outlier cases and thus were missing from the SuperPRO sample. The third subsection describes how we combined data from both of our sources to adjust our decomposition of the CMI change to include an estimate for outlier cases. The final subsection describes the generalized least squares (GLS) algorithm we used to reconcile the SuperPRO estimate of the components of change with the total change found on the MEDPAR file.

SUPERPRO SAMPLE

SuperPRO Review Process

The SuperPRO is charged with reviewing the work of the PROs. The PROs monitor the quality of inpatient care, the necessity of admissions, and the accuracy of DRG assignments. They select cases to review for a large variety of reasons and review the complete medical record for each selected case. The SuperPRO selects a sample of each PRO's cases and obtains copies of the medical records from the PRO. Assisted by physicians, highly trained medical record technicians at the SuperPRO reabstract each of the medical records, assigning a primary diagnosis and applicable secondary diagnoses, and recode all procedures described in the medical record. These new diagnoses are fed into the official Grouper program to obtain the DRG associated with these patient characteristics.

The cases reviewed by the SuperPRO are selected randomly from all the cases reviewed by the PROs. The only cases used in our analysis were those that were part of the 3 percent random sample, which is a sample of all cases in each PRO area, excluding outlier cases. All other cases reviewed by the SuperPRO have idiosyncratic sampling histories that make it impossible to use their data to infer population rates. The SuperPRO manually exam-

ined the PRO worksheet for each case to select cases for the random sample.

The SuperPRO organizes its review activities around the *cycles* of each PRO's contract year. FY 1988 data are found predominantly in cycles 7 and 8. Cycle 7 consists of discharges reviewed by the PRO from either January 1 through June 30, 1988, or February 1 through July 29, 1988. Cycle 8 consists of discharges reviewed from August 1, 1988, through January 31, 1989. In general, earlier cycles cover earlier six-month intervals.[1]

Although the cycles are defined in terms of month of PRO review, they are strongly correlated with month of discharge. In 86 percent of the FY 1988 cases reviewed in cycle 7, the patient was discharged in the first six months of FY 1988. In 93 percent of the cases reviewed in cycle 8, discharge occurred in the last six months of FY 1988.

Description of SuperPRO Data

For the random sample cases reviewed by the SuperPRO, we obtained the diagnostic and procedure codes assigned by each of three agencies—the hospital, the PRO, and the SuperPRO. The SuperPRO codes were assigned during the SuperPRO's regular review of cycle 7 and 8 cases during 1989. We then ran the 1987 and 1988 Groupers on each case to determine the effect of each agency's coding on DRG assignment.[2]

We also obtained data from a reabstraction sample, which was a subsample of the FY 1987 cases that the SuperPRO recoded using its 1988 coding practices. The SuperPRO began its recoding in October 1989 and completed it by the year's end, which means the reabstraction of the 1987 cases and the abstraction of the 1988 discharges were performed within the same year. Thus, we believe that the coding standards used for the reabstraction were essentially similar to those used for the 1988 discharges. The October 1, 1987, list of codes was used for both abstractions.

---

[1]Carter, Newhouse, and Relles (1990) give information about the timing for cycles 3 through 6.

[2]Diagnostic and procedure codes were translated to match the Grouper in use.

Sample Sizes

The discharges from FY 1987 and FY 1988 spanned cycles 4 through 8 (see Table 2.1). Data for all PPS states were available for cycles 4 through 6. However, in cycles 7 and 8, the SuperPRO reviewed PROs in only 23 states. Thus, our decomposition of change was restricted to these 23 states. We used data on the other states to compare coding behavior for 1987 discharges in the 23 sample states to that for the nation as a whole.

Table 2.2 lists the 23 states for which either cycle 7 or 8 data were available. These states cover all regions of the United States and account for approximately 56 percent of the country's Medicare discharges. Column 3, total SuperPRO sample, includes all FY 1988 cases, including those from cycle 6. The next two columns show the number of sample cases from each of cycles 7 and 8. Many of the states were not reviewed during cycle 8, and four states were not reviewed during cycle 7. For states for which we had only one cycle of data, we had substantial sampling variation by discharge quarter.

Our sample is effectively a stratified sample whose sampling fractions are unequal and whose strata are defined by state and discharge quarter. For some analyses, we had to calculate weighted statistics. For example, to measure real change, we compared the CMI for the FY 1988 discharges in cycles 7 and 8 to the CMI for

Table 2.1

Size of SuperPRO Random Sample by Cycle and Fiscal Year of Discharge

	Number of Cases			
	All States		Sample States	
Cycle	FY 1987	FY 1988	FY 1987	FY 1988
4	385	0	172	0
5	1775	0	857	0
6	1958	375	1069	90
7	N/A	N/A	86	879
8	N/A	N/A	5	505
Total	4118	375	2189	1474

NOTE: Only cases for which the diagnostic and procedure codes assigned by the hospital, PRO, and SuperPRO were valid are included. Data from Puerto Rico are excluded.

Table 2.2

FY 1988 Sample by State

State	Discharges[a] (1000)	Total SuperPRO Sample	Cycle 7	Cycle 8	Real Change Sample
AL	39	82	37	44	81
AR	27	44	41	0	41
CO	16	43	43	0	39
CT	23	43	0	35	34
HI	5	41	0	30	28
IL	98	62	55	0	55
IA	26	90	50	39	89
KS	24	46	46	0	46
MA	51	175	48	127	175
MS	31	85	78	0	78
NH	7	23	23	0	18
NY	137	81	80	0	75
NC	47	86	40	46	86
ND	7	92	47	40	87
OH	96	58	51	0	51
OR	20	39	36	0	36
PA	124	44	37	0	6
TN	51	50	50	0	45
TX	108	71	0	60	58
UT	8	29	29	0	29
VA	42	60	0	43	36
WV	21	49	49	0	49
WI	42	81	39	41	80
Total	1048	1474	879	505	1322

[a]Twenty percent sample from MEDPAR file.

the reabstracted sample of FY 1987 discharges in cycles 5 and 6. We restricted this analysis to state and discharge-quarter pairs that contained more than five sample discharges. Thus, the real change sample given in the last column of Table 2.2 is the sum of the cycle 7 and 8 sample minus the few discharges that occurred in quarters with five or fewer discharges.

Table 2.3 describes the FY 1987 discharges in each sample state. The column listing total SuperPRO samples counts discharges from cycles 4, 5, and 6. We used these discharges to compare hospital coding practices relative to the SuperPRO's coding practices with similar data for the FY 1988 discharges in cycles 6, 7, and 8. The next column describes the reabstraction sample in the sample states, which consisted of essentially all cycle 5 and 6 cases in those states. The SuperPRO reabstracted a total of 3600

Table 2.3

FY 1987 Sample by State

State	Discharges[a] (1000)	Total SuperPRO Sample	Reabstraction Sample	Real Change Sample
AL	41	86	85	85
AR	28	66	61	45
CO	16	63	63	31
CT	21	120	97	59
HI	5	155	155	94
IL	96	123	86	70
IA	27	96	96	96
KS	24	81	79	76
MA	51	79	79	79
MS	30	155	153	102
NH	8	59	59	20
NY	135	88	88	51
NC	45	76	71	71
ND	7	74	73	73
OH	97	100	74	58
OR	21	56	55	25
PA	122	56	41	10
TN	50	85	85	44
TX	106	91	74	45
UT	7	63	62	50
VA	44	150	107	57
WV	21	83	82	64
WI	40	93	93	93
Total	1041	2098	1918	1398

[a]Twenty percent sample from MEDPAR file.

cases from FY 1987, including 1682 in states not in our cycle 7 and 8 sample states, so that we could describe how changes in coding standards affected the national CMI. The final column of Table 2.3 gives the FY 1987 sample sizes used to measure real change. The numbers in this column are equivalent to the numbers of cases in the reabstraction sample after excluding all cases from quarters in which the matching FY 1988 quarters contained five or fewer cases.

Treatment of Stratified Sample

As one can see from the data presented in Tables 2.2 and 2.3, the sample contains widely varying proportions of the population in each state. When the parameter being estimated varies from

state to state or from quarter to quarter, such as the CMI does, the mean of the sample will be a biased estimate of the population. The usual solution to this problem is to calculate a weighted sample mean with the weights on each sample data point being proportional to the ratio of the population size to the sample size for the strata (state and discharge quarter, in our case) containing the sample. However, when the sampling proportions are as varied as they are in our case, the weighted sample mean is very imprecise because of the large weights given to just a few data points.

Our method of dealing with the stratification issue was to first examine whether we could establish the existence of state differences in the statistic in question. We did this using an analysis of variance (ANOVA) in which the explanatory variables included dummy variables for each state. If the collection of state effects was found to be significant, we would use the weighted mean. This turned out to be the situation with measurement of true change. If, on the other hand, the state effects were statistically indistinguishable from 0, they might still exist and be small. So, if there was an a priori reason to believe that state effects might exist, as in the case of hospital coding, we would also calculate the weighted mean in this case. If the weighted mean then turned out to be in the same neighborhood as the unweighted mean, we would feel confident using the unweighted mean, which would be more precise than the weighted mean.

MEDPAR DATA

To estimate the CMI level at which Medicare paid for its cases—i.e., the *paid CMI*—during FY 1987 and FY 1988, we used a 20 percent sample of Medicare cases chosen from the MEDPAR file on the basis of the last digit of the beneficiary's identification number. We also used this sample file to develop sampling weights to go from sample statistics to estimates for all cases in our 23-state sample. Sample sizes are given in Table 2.4. The sample states contained 56.52 percent of discharges during FY 1987 and 56.58 percent of discharges during FY 1988.

In addition, we used the MEDPAR file to estimate the frequency of characteristics of outlier cases correlated with changes in coding behavior, because outlier cases are omitted from the random sample cases reviewed by the SuperPRO. Finally, we used the in-

Table 2.4

MEDPAR Sample Sizes

State Group	Discharges Total	Outlier
All states, FY 1987	1,842,108	79,211
All states, FY 1988	1,851,951	88,760
Sample states, FY 1987	1,041,120	48,920
Sample states, FY 1988	1,047,883	54,872

formation about total CMI change found in the MEDPAR data to adjust our estimates of the components of CMI change from the SuperPRO sample. These two uses of the MEDPAR data are discussed, respectively, in the next two subsections.

ADJUSTMENT FOR OUTLIER CASES

The SuperPRO data base does not contain outlier cases. Outliers have much higher weights than other cases, and the fraction of cases that were outliers increased from 4.7 percent in FY 1987 to 5.2 percent in FY 1988 in our sample states.[3] Thus, it was necessary to adjust our estimates to account for the missing outlier cases.

The adjustment was based on what we learned about the difference between SuperPRO and hospital coding. We used the sample of nonoutlier cases at the SuperPRO to build regression models that estimate SuperPRO coding as a function of characteristics of the DRG assigned by the hospital and PRO. Then we used the DRG characteristics of outlier cases to estimate how the SuperPRO would have coded outlier cases.

The dependent variable for one regression was the weight assigned by the SuperPRO at its original coding minus the weight assigned by the hospital. The dependent variable for another regression was the weight assigned by the original SuperPRO coding minus the weight assigned by the PRO. The final two dependent variables were the weight assigned by the SuperPRO during recoding minus the weight assigned by, in one case, the hospital and, in the other case, the PRO. Separate regressions were run

[3]In the nation as a whole, the fraction of cases that were outliers increased from 4.3 percent in FY 1987 to 4.8 percent in FY 1988. The number of cases is shown in Table 2.4.

for each fiscal year using the dependent variables based on the SuperPRO's original coding, yielding a total of six regressions. All DRG weights in each regression were based on the 1988 Grouper.[4]

The explanatory variables of each regression were characteristics of the DRG assigned by the hospital or PRO as appropriate to the dependent variable. For example, one of the coding changes we observed was additional coding of complicating conditions. One of the explanatory variables for the regression was the increase in DRG weight that occurs if a complicating condition is coded. These conditions affect DRG weight only for the uncomplicated DRG within a DRG pair, so other DRGs have a value of 0 for this explanatory variable.

We used the MEDPAR file to calculate the mean value of each explanatory variable for the set of outlier cases in each year. We then put these values into the estimated regression equations to get predictions. The regression predictions based on the hospital and PRO codings were averaged with a weight on the hospital prediction of 0.75 and a weight on the PRO prediction of 0.25. These weights were chosen because PROs review approximately 25 percent of cases, and when PROs disagree with the DRG assigned by the hospital, the PRO's DRG prevails. For the regression based on the original SuperPRO coding, we got an estimate of the difference between how the SuperPRO would have coded outlier cases had it coded them during its regular PRO review each year and how those cases were actually coded for payment each year. For the other regression, we got an estimate of the difference between how the SuperPRO would have coded the 1987 outlier cases using 1988 standards and how they were coded for payment in 1987. From these estimates of coding differences, we predicted the CMIs that would have resulted had the SuperPRO assigned DRGs to outliers during each of its codings. These predictions assume that outlier cases are coded the same way as cases with the same characteristics that are not outliers.

The next step of the adjustment was to average the CMIs for outlier and nonoutlier cases in each year using weights equal to their respective population proportions on the MEDPAR file. For each of the SuperPRO CMIs, the values for nonoutlier cases were from

[4]We also ran all six regressions using the 1987 Grouper weights. Although we do not report the details of those regressions, they were used to produce the data in App. A.

the SuperPRO sample and the values for outlier cases were from the regression predictions. For the hospital coding, the nonoutlier cases were from the SuperPRO sample and the outlier cases were from the MEDPAR data. Using the estimated overall CMIs (including outlier cases), the components of the increase were calculated as described in Sec. 1.

ESTIMATING THE COMPONENTS OF CMI CHANGE

We computed the components of the CMI change from the SuperPRO sample by determining case-level estimates of those components, then computing (weighted) means over the SuperPRO sample, and then adjusting the means for the omission of outlier cases. These components add up to total change—i.e., an unbiased estimate of the total CMI change found in the MEDPAR data for the sample states. Because we had only a sample of cases, the sample mean differs from the more precise population mean estimate from MEDPAR. We exploited the more precise estimate to simultaneously adjust our estimates of the components of change according to the postulated regression model described next.

Consider the following regression problem. Let

μ_1 = change in the paid CMI from 1987 to 1988,

μ_2 = true change in the CMI from 1987 to 1988,

μ_3 = coding change from 1987 to 1988 resulting from changes in implicit or explicit standards (that is, coding standard changes common to SuperPRO and hospitals),

μ_4 = coding change from 1987 to 1988 resulting from hospital practices only,

μ_5 = Grouper normalization effect,

y_1 = estimate of μ_1 derived from the MEDPAR sample,

y_2 = estimate of μ_2 derived from the SuperPRO sample,

y_3 = estimate of μ_3 derived from the SuperPRO sample,

y_4 = estimate of μ_4 derived from the SuperPRO sample,

y_5 = estimate of μ_5 derived from the MEDPAR sample,

where the μ's correspond, respectively, to Eqs. 1 through 5 in Sec. 1. Define ε_1 as the error term in the equation $y_1 = \mu_1 + \varepsilon_1$. Because by definition $\mu_1 = \mu_2 + \mu_3 + \mu_4 + \mu_5$, we have

$$y_1 = \mu_2 + \mu_3 + \mu_4 + \mu_5 + \varepsilon_1.$$

Also,

$$y_2 = \mu_2 + \varepsilon_2,$$

$$y_3 = \mu_3 + \varepsilon_3,$$

$$y_4 = \mu_4 + \varepsilon_4,$$

$$y_5 = \mu_5 + \varepsilon_5,$$

where $E(\varepsilon_i) = 0$, and ε_2, ε_3, and ε_4 are each independent of both ε_1 and ε_5 and have a covariance matrix that we can estimate from the SuperPRO data.

If the ε_i's were all independent, this would be a standard weighted regression problem, with y_i, the weight on each of the five data points, inversely proportional to the estimated variance of ε_i. But the ε_i's are not independent; thus, we have a nondiagonal weight matrix and so must use GLS.

Formally, let

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix},$$

and define W to be the covariance matrix of ε. In matrix terms, our formulation is

$$y = X\mu + \varepsilon,$$

where ε has mean 0 and an estimated covariance matrix W. According to the well-known principles of GLS (see Wonnacott

and Wonnacott, 1979), the minimum variance, linear, unbiased estimate of μ is

$$\mu = (X'W^{-1}X)^{-1}\ (X'W^{-1}y),$$

which has a covariance matrix

$$\mathrm{var}(\mu) = (X'W^{-1}X)^{-1}.$$

Section 3, which describes our empirical results, includes both the initial and the GLS-adjusted estimates of μ.

3. RESULTS

We begin our discussion of findings with an examination of trends in the consistency with which the three types of agencies—hospitals, PROs, and the SuperPRO—code the same cases. We then look at the CMI's growth from 1987 to 1988, first showing the measured increase and then decomposing it into (1) the increase caused by true change; (2) the increase caused by changes in hospital coding practices relative to the SuperPRO coding standard (i.e., hospital-only changes); (3) the increase caused by changes in the coding standard (i.e., the hospital and SuperPRO changes); and (4) the increase caused by the Grouper normalization effect. We then describe the adjustment we made to account for the missing outlier cases, after which we show how we developed final estimates for our 23 states and then provide estimates of the components of change in the national CMI.

TRENDS IN CODING AGREEMENT

Table 3.1 compares the DRGs assigned by the hospitals, PROs, and SuperPRO. The first section of the table shows the percentage of cases in which the SuperPRO, during its normal PRO review, assigned the same DRG the hospital had assigned at discharge. Although the 1985 data are quite different from data for the other years,[1] the increase in the rate of agreement in recent years has been tiny, with the two codings agreeing in roughly 85 percent of the cases. Somewhat surprisingly, since 1986 the SuperPRO has assigned a lower DRG weight than the hospital in only 0.5 percent more cases than it has assigned a higher DRG weight.

The middle portion of Table 3.1 compares PRO and SuperPRO coding. The findings here are generally similar to those for hospital versus SuperPRO coding, although there is slightly greater agreement. The bottom section of Table 3.1 shows that the PROs only rarely disagree with the DRG assigned by the hospital.

[1]Chi-square tests show a lack of independence of year and agreement category for all three sections of Table 3.1. The chi-square statistics, each with six degrees of freedom, are 50.4 ($p < .001$), 26.9 ($p < .001$), and 14.1 ($p = .03$).

Table 3.1

Comparison of Weights of DRGs Assigned by Different Coders, Current-Year Coding, All States (in percent)

	Hospital vs. SuperPRO			
FY	Hospital Higher	Same	SuperPRO Higher	Total
1985	15.4	76.9	7.7	100
1986	8.5	83.5	8.0	100
1987	8.1	84.4	7.5	100
1988	7.2	86.0	6.8	100
	PRO vs. SuperPRO			
FY	PRO Higher	Same	SuperPRO Higher	Total
1985	11.1	82.4	6.5	100
1986	6.8	86.6	6.6	100
1987	6.8	87.2	6.0	100
1988	5.4	89.8	4.8	100
	Hospital vs. PRO			
FY	Hospital Higher	Same	PRO Higher	Total
1985	5.3	91.5	3.3	100
1986	3.1	94.0	2.9	100
1987	2.8	94.6	2.6	100
1988	3.0	94.0	3.0	100

NOTE: Data for 1988 are predominantly from 23 sample states rather than all states.

Table 3.2 also compares the weights of assigned DRGs, but only for the 23 sample states that had a PRO review by the SuperPRO during cycle 7 or 8. As can be seen, the percentage of cases in each agreement category for the sample states is very similar to that for the nation as a whole in 1985 through 1987. It is also very similar for 1988, but, as noted in Table 3.1, the 1988 "national" data came predominantly (89 percent) from the sample states, which would, of itself, cause the statistics for the two samples to be very similar. Thus, although we cannot be sure that the 1988 data are nationally representative, the fact that the sample states showed nationally representative coding behavior from 1985 to 1987 increases the likelihood that they are.

Table 3.2

Comparison of Weights of DRGs Assigned by Different Coders, Current-Year Coding, Sample States (in percent)

	Hospital vs. SuperPRO			
FY	Hospital Higher	Same	SuperPRO Higher	Total
1985	15.6	76.1	8.2	100
1986	7.8	84.7	8.5	100
1987	8.3	84.3	7.5	100
1988	7.0	86.2	6.9	100
	PRO vs. SuperPRO			
FY	PRO Higher	Same	SuperPRO Higher	Total
1985	10.1	83.7	6.2	100
1986	6.2	87.3	6.5	100
1987	6.6	88.1	5.3	100
1988	5.2	90.4	4.5	100
	Hospital vs. PRO			
FY	Hospital Higher	Same	PRO Higher	Total
1985	6.2	89.6	4.2	100
1986	3.1	93.4	3.5	100
1987	3.2	93.5	3.3	100
1988	3.1	93.5	3.4	100

DECOMPOSITION

Table 3.3 shows the CMIs for the sample states and the nation in 1987 and 1988 as measured by the Grouper used in each year (i.e., the paid CMI). The sample states experienced a CMI increase of only 3.0 percent compared to a national increase of 3.3 percent.

True Change

Table 3.4 shows our estimates of true change from 1987 to 1988 for nonoutlier cases in the sample states. The estimated CMI for FY 1987 comes from the SuperPRO's recoding of the 1987 cases using the same coding practices it used for the FY 1988 cases. The measurement is very imprecise, i.e., the standard errors are relatively high.

Table 3.3

Measured Change in CMI

	Paid CMI			Percent
	FY 1987	FY 1988	CMI Increase	Increase
Sample states	1.23082	1.26813	0.03731	3.0
All PPS states	1.24081	1.28127	0.04046	3.3

Table 3.4

True Change in CMI for Sample States (Excluding Outlier Cases)

	CMI		Change	
	FY 1987	FY 1988	Average	Standard Error
1988 Grouper	1.25560	1.27283	0.01723	0.04467

NOTE: Weights based on state and discharge quarter were used in calculating CMI.

Hospital-Only Coding Changes

We next measured the component of the CMI increase resulting from coding practice changes that occurred at hospitals but not at the SuperPRO during the same period. This hospital-only coding component of the CMI increase was calculated by comparing the DRG weight assigned by the SuperPRO's original coding to the DRG weight that was paid. PROs review about one-quarter of all discharges and determine the DRG that is paid for those cases (ProPAC, 1989, p. 33). Thus, we measured the hospital-only coding effect as a weighted average of the change in the difference in the CMI assigned by hospitals and the SuperPRO (weight = 0.75) and the change in the difference in the CMI assigned by the PROs and the SuperPRO (weight = 0.25).[2]

Table 3.5 shows the calculations. For FY 1987 discharges, the SuperPRO's coding resulted in 1988 Grouper DRGs with significantly higher weights than did the coding of the hospitals and PROs. This

[2]The results are insensitive to the values of these weights because of the similarity of hospital and PRO coding. Table B.1 in App. B presents data separately for the hospital and PRO.

Table 3.5

Effect of Change in Hospital Coding Relative to SuperPRO Standard on Change in CMI

Sample	CMI		Increase over SuperPRO	
	SuperPRO	Hospital/PRO	Average	Standard Error
Sample states				
FY 1987	1.20372	1.19116	–0.01256[a]	.0055
FY 1988	1.26413	1.25970	–0.00443	.0075
Difference			0.00813	.0095
All states				
FY 1987	1.19460	1.18043	–0.01417[a]	.0041

[a] $p < .05$.

NOTE: Each case has equal weight in calculating CMI.

finding is true for both the sample states and the PPS states. Most of this difference had disappeared by 1988, when the 1988 Grouper was in use.

We examined the DRGs assigned by the hospitals, PROs, and SuperPRO. Although many DRG disagreements were difficult to characterize, two patterns were clear. First, the SuperPRO frequently lowered the weight for cases to which the hospitals or PROs had assigned DRG 468, which is used when the patient's operating room procedure is not related to his principal diagnosis and which is assigned in over 1 percent of Medicare cases.[3] In addition, in 1987 the SuperPRO was more likely than the hospitals to note complications and comorbidities. Using the 1987 Grouper, the effects essentially cancelled each other out, resulting in the data of Table 3.2, in which it can be seen that cases were equally likely to be assigned to higher- and lower-weighted DRGs. However, the 1988 Grouper put more weight on the presence of complications and comorbidities, thereby assigning a significantly higher weight to the SuperPRO coding of 1987 cases than to the hospital and PRO coding.

As discussed in Sec. 2, our sample contained different proportions of the cases from each sample state and from each month of the year. Our previous research found state and discharge month to have no effect on the difference between SuperPRO and hospital/PRO coding (Carter, Newhouse, and Relles, 1990). We used ANOVA to test

[3]The SuperPRO was almost equally likely to change the principal diagnosis and the procedure code.

whether state and/or discharge month effects were present in the cycle 7 and 8 data, and found that we could not reject the hypothesis of no effects.[4] However, since this test may lack power and state effects may exist, we also calculated each agency's CMI using weights proportional to the ratio of MEDPAR discharges to SuperPRO sample cases in each combination of state and discharge quarter. Table 3.6 compares the unweighted estimate of the hospital-only component of the CMI change. The unweighted effects were taken directly from Table 3.5, and the weighted numbers were calculated analogously after weighting by state and discharge quarter. The weighted effect is 0.00362 higher; thus, our choice of an unweighted rather than a weighted estimate affects roughly one-tenth of the total CMI increase to be decomposed.

Changes in Both Hospital and SuperPRO Coding

There were measurable differences between the SuperPRO's original coding of the 1987 cases and its recoding of them using FY 1988 coding rules. The average DRG weights assigned using each coding are shown in Table 3.7. For these 1987 cases, SuperPRO data were available for all states and were recoded using 1988 coding rules. The sample size for the 23-state sample was 58 percent of the all-state sample. The discharges from the sample states received somewhat less upcoding than those from the national sample. The changes in coding practice that occurred at the SuperPRO almost certainly occurred at hospitals also, since the 1988 coding of the two agencies was so similar.

Table 3.6

Weighted and Unweighted Estimates of Effect of Changes in Hospital Coding Relative to SuperPRO Standard on Change in CMI

	Effect	
	Average	Standard Error
Unweighted	0.00812	0.0095
Weighted	0.01174	0.0135

[4]For state, F = 0.90 with 22 and 1360 degrees of freedom (p = .60); for discharge month, F also = 0.90 with 11 and 1371 degrees of freedom (p = .54).

Table 3.7

Effect of Changes in Hospital and SuperPRO Coding on Change in CMI (1987 Cases)

	CMI		Change	
	Original Coding	New Coding	Average	Standard Error
Sample states	1.20392	1.21307	0.00915[a]	0.00488
All states	1.19476	1.20802	0.01326[b]	0.00353

[a] $p = .06$.

[b] $p < .001$.

NOTE: Each case has equal weight in calculating CMI.

The hospital-and-SuperPRO-coding component of the CMI increase is given as the average change in Table 3.7. Each case was given equal weight in calculating this component. There is no reason to believe that SuperPRO coding is affected by state or month of discharge.

Two kinds of cases were responsible for 93 percent of the increase in the CMI measured by the 1988 Grouper from the first coding to the recoding.[5] Ventilator support, recorded 11 times in the recoding, had been recorded in only one of those 11 cases the first time. In the first coding, which was performed to verify the accuracy of PRO DRG assignment using the 1987 Grouper, ventilator support had no effect on DRG accuracy. Ventilator support causes assignment of DRG 475, with its high weight of 3.1757, under the 1988 Grouper. The ventilator support cases were responsible for over one-third of the increase in the 1988 Grouper CMI for the recoding over the original coding. The MEDPAR data recorded an increase from 1987 to 1988 in the frequency of ventilator support that exactly corresponded to this measured increase in coding behavior. In 1987, ventilator support was noted in only 0.04 percent of cases. In 1988, ventilator support was noted roughly ten times as frequently—in 0.39 percent of cases.

The second kind of case responsible for the CMI increase in the recoding was paired DRGs, which accounted for 56 percent of the 3600 cases that were recoded nationally. Excluding ventilator support cases, the CMI increase in these paired DRGs averaged six times the increase in nonpaired, nonventilator cases.

[5]We had hypothesized that we would also find an increase in the recording of tracheostomies (DRG 474), but this procedure is so rare that not one was observed in the SuperPRO sample. In the MEDPAR file, DRG 474, as assigned by the 1988 Grouper, increased from 0.09 percent of FY 1987 cases to 0.13 percent of FY 1988 cases.

Grouper Normalization Effect

The Grouper normalization factor is the difference between the 1987 CMI measured with the 1988 Grouper and the 1987 CMI measured with the 1987 Grouper. It essentially measures misnormalization of the 1988 Grouper and is part of the effect of coding and administrative practices.

Table 3.8 gives the CMI measured by each Grouper on the MEDPAR file. The first row covers cases in the sample states in FY 1987. The 1987 and 1988 Groupers measure almost the same CMI. The difference between the two Grouper measurements (last column in the table) is less than 0.1 percent of the CMI. This difference is the Grouper normalization effect used in analyzing the components of change measured by the 1988 Grouper. Its small size shows that the misnorming of the 1988 Grouper had only a small effect on the CMI increase in the sample states and even less effect on the increase in the national CMI.

ADJUSTMENT FOR OUTLIER CASES

To complete our decomposition, we needed to adjust the estimates of true change and of coding changes to account for the outlier cases that were missing in the SuperPRO data base. Our adjustment was based on what we knew about the differences between SuperPRO and hospital coding.

Table 3.8

CMI by Grouper and Fiscal Year

Set of Discharges	CMI 1987 Grouper	CMI 1988 Grouper	Grouper Normalization Component
Sample states			
FY 1987	1.23082	1.23394	0.00312
FY 1988	1.25361	1.26813	
All PPS cases[a]			
FY 1987	1.24081	1.24175	0.00094
FY 1988	1.26751	1.28127	

[a]Excludes Puerto Rico.

NOTE: These data are based on a 20 percent sample of the MEDPAR file.

For each of the two SuperPRO codings, we built a regression model that estimated coding differences as a function of DRG characteristics, such as being the uncomplicated DRG within a pair of DRGs. Then we used the MEDPAR file to determine how many outlier cases had each characteristic. When the outlier characteristics were placed in the original SuperPRO coding regression, we got a prediction of the coding effect and thus of the CMI that would have resulted if the SuperPRO had coded outliers during its regular PRO review. For the 1987 cases, we used a second regression to get an estimate of how the SuperPRO would have coded the outlier cases using its FY 1988 standards.

Table 3.9 shows the regressions of the SuperPRO weight from the original coding minus the weight assigned by the hospital.[6] Separate regressions were fit for each Grouper and fiscal year. The data used to fit the regressions were restricted to the sample states, although the 1987 national data gave very similar results. Only a small

Table 3.9

Regression of Weight Assigned by SuperPRO's Original Coding Minus Weight Assigned by Hospital

Hospital Coding Variables	FY 1987		FY 1988	
	Coefficient	t-Statistic	Coefficient	t-Statistic
Increase for complication or comorbidity	0.2374	9.81	0.2323	7.04
DRG 468	–0.2163	3.64	–0.3696	4.89
Surgical DRG	–0.0378	1.37	–0.0067	0.18
DRG weight	–0.0962	4.98	–0.0834	3.24
Weight for surgical cases	0.0713	3.35	0.0453	1.62
Intercept	0.0822	4.27	0.0706	2.71
R^2	0.09		0.09	

[6]Table B.2 in App. B provides information about the regression of the difference between the SuperPRO and PRO coding. The results there are generally similar to those here, except that in FY 1988 only, the PROs' coding of DRG 468 evidently resembled that of the SuperPRO more than that of the hospitals.

percentage of the variance is explained, but the t-statistics show that the coefficients are reasonably precise.[7]

The hospital coding variables listed in the table describe the hospital's coding of each case. The first explanatory variable is the increase in weight that occurs if an additional complicating condition or comorbidity is coded. This variable is 0 for cases that are not in paired DRGs and for cases for which the hospital coded a complication or comorbidity. As expected, the coefficient for this variable is quite large and positive. The second explanatory variable is just a dummy variable; it is 1 for cases to which the hospital assigned DRG 468 and 0 otherwise. Cases to which the hospital assigned DRG 468 were given lower weights on average by the SuperPRO.

The remaining explanatory variables are the same ones used in our study of the CMI increase from 1986 to 1987: a dummy variable indicating surgical DRGs, the weight assigned to the case, and the interaction of the DRG weight and surgical dummy. The coefficients show that the SuperPRO decreased weights more for high-weight cases than for low-weight cases and that this relationship was much stronger for medical than for surgical cases. The surgical dummy is not significantly different from 0, but is kept in the regression to get unbiased estimates of the differential effects of DRG weight for medical and surgical cases.

Table 3.10 contains a similar regression of the difference between the SuperPRO's recoding of the 1987 cases and the hospital's original coding of those cases. An additional explanatory variable was added to the regression; it is 1 if the hospital assigned a DRG in MDC 4 to the case and did not note ventilator support. These are the only cases that can be upcoded into DRG 475. Although the variable is not significant in the sample states, it is significant in the set of all states, and we believe it is a reasonable predictor of actual coding change.

We used the MEDPAR file to calculate the mean value in the set of each year's outlier cases for each explanatory variable in the regressions contained in Tables 3.9 and 3.10. These mean values were then plugged into the regression equations to estimate how the SuperPRO would have coded outlier cases and the resulting CMIs. Table B.4 in App. B provides details on the CMI calculations.

---

[7]We cannot rule out the possibility that the coefficients suffer from omitted-variable bias. However, we believe we understand the coding process and have selected the most important DRG characteristics for inclusion in the model.

Table 3.10

Regression of Weight Assigned by SuperPRO's New Coding Minus Weight Assigned by Hospital (FY 1987 Cases)

Hospital Coding Variables	Sample Sates		All PPS States	
	Coefficient	t-Statistic	Coefficient	t-Statistic
Increase for complication or comorbidity	0.1957	6.30	0.2276	9.95
DRG 468	–0.2808	3.68	–0.1599	2.65
MDC 4, no ventilator	0.0267	1.19	0.0503	3.05
Surgical DRG	–0.0160	0.45	–0.0313	1.18
DRG weight	–0.1006	3.79	–0.1065	5.45
Weight for surgical cases	0.0522	1.82	0.0639	3.00
Intercept	0.1024	4.06	0.1057	5.70
R^2	0.07		0.06	

Table 3.11 shows the original estimates of the components of change from the nonoutlier data and how these estimates change with the inclusion of outlier cases. Since outliers constituted only 5 percent of cases, the effects were generally modest—each component estimate changed by less than 10 percent in the sample states. All of the components decreased in size, which is consistent with our expectations given that the CMI for outlier cases measured by the 1988 Grouper decreased from 1987 to 1988.

FINAL ESTIMATES

Sample State Estimates

After deriving estimates of each component of the CMI increase as described above, we modified them to take into account the CMI change measured for the sample states from the MEDPAR file. The estimates we derived for each component are repeated in the first column of Table 3.12. The Grouper normalization component is from Table 3.8; the other components are from Table 3.11. Because the SuperPRO data contained only a sample of cases, the sum of the components of change do not add to the actual change measured on the MEDPAR file. Rather, the components underestimate total change by 6 percent.

We used the GLS methodology described in Sec. 2 to reconcile these estimates to the actual change. The last two columns of Table 3.12

Table 3.11

Estimated Effect of Including Outlier Cases in SuperPRO Data Base on Components of CMI Increase

Component[a]	Size of Effect: Excluding Outliers	Including Outliers
Sample states		
True change	0.01723	0.01567
Hospital-only coding changes	0.00813	0.00756
Hospital and SuperPRO coding changes	0.00916	0.00862
All states		
Hospital and SuperPRO coding changes	0.01326	0.01296

[a]The Grouper normalization component is not listed because it already includes outliers.

Table 3.12

Original and GLS Estimates of Components of CMI Change in Sample States

Component	Original Estimate: CMI Increase	Original Estimate: Standard Error	GLS Estimate: CMI Increase	GLS Estimate: Standard Error
Measured change	0.03731	0.0012	0.03731	0.0012
True change	0.01567	0.0425	0.01798	0.0095
Coding/administrative changes				
Hospital only	0.00756	0.0086	0.00758	0.0085
Hospital and SuperPRO	0.00862	0.0047	0.00863	0.0047
Grouper normalization	0.00312	0.0004	0.00312	0.0004
Subtotal	(0.01930)	0.0094	(0.01933)	0.0094

describe the GLS estimates of the components, which do sum to the measured change. In the original estimates, the true change component had by far the largest standard error (reflecting both the weighting of this component and the fact that the coding components have smaller standard deviations per case). The GLS procedure recognized this higher variability, so most of the adjustment required to enable the components to sum to the total change was made to the true change component.

National Estimates

We made national estimates of the effects of two of the components of CMI change: coding changes common to both the hospital and the SuperPRO, and the Grouper normalization factor. In doing our analysis of changes in hospital coding relative to the SuperPRO standard, we did not find evidence of state-level effects. Thus, it is not unreasonable to assume that the estimate of the hospital-only coding component of CMI change that we derived from the 23 sample states is also a reasonable estimate for all states. If we accept this assumption, then we can estimate true change in the national CMI as a residual.

Table 3.13 provides the national decomposition. The results are quite similar to the results for the sample states.

Table 3.13

Decomposition of National CMI Increase

Component	CMI Increase from Component
Measured change	0.04046
True change	0.01898
Coding/administrative changes	
Hospital only	0.00758
Hospital and SuperPRO	0.01296
Grouper normalization	0.00094
Subtotal	(0.02148)

4. CONCLUSIONS

Table 4.1 is our best estimate of the decomposition of the 3.0 percent increase in CMI in the sample states. To produce this summarizing table, we expressed the CMI point increases estimated in Table 3.12 as both a percentage of the total CMI increase and an annual rate of increase.

Our point estimate is that true change increased the CMI at an annual rate of 1.5 percent. Because the small size of our sample limits the precision of our estimates, however, we can say only that there is an 80 percent chance that the rate of true change in the sample states was between 0.0 and 2.4 percent. Nevertheless, we can be sure that the coding and Grouper changes contributed to the rate of increase in the CMI between 1987 and 1988. An 80 percent confidence limit on the part of the 1987–1988 increase in the CMI that was *not* caused by changes in patient resource needs is 0.6 to 3.0 percent. Thus, a substantial fraction of the increase was due to factors that the HCFA would not ideally wish to reimburse.

The all-states data in the table rely on the assumption that all hospitals changed their average coding relative to the SuperPRO coding standard in the same way that the sample states did. Although we cannot be sure about what happened in FY 1988, we do know that FY

Table 4.1

Summary of Decomposition of CMI Increase (in percent)

	Sample States		All States	
Component	Portion of Total CMI Increase	Annual Rate of Increase	Portion of Total CMI Increase	Annual Rate of Increase
True change	48	1.5	47	1.5
Coding/administrative changes				
Hospital only	20	0.6	19	0.6
Hospital and SuperPRO	23	0.7	32	1.0
Grouper normalization	8	0.3	2	0.1
Subtotal	(52)	(1.6)	(53)	(1.7)
Total	100	3.0	100	3.3

1987 coding behavior relative to the SuperPRO coding standard was similar for the hospitals and PROs in our 23-state sample and in the nation as a whole. We also know that from 1985 through 1987, trends in the rate of agreement between the SuperPRO and the hospitals and PROs were similar in our sample states and in the nation as a whole. Thus, it is quite plausible that the estimate of coding behavior relative to the SuperPRO standard from the sample also fits the nation as a whole. The all-states data are merely a transformation of the data from Table 3.13.

Our estimates assume that medical records did not become more complete between 1987 and 1988. We believe this assumption to be approximately correct.[1] To the extent that it is incorrect, less of the measured change is the result of true change. Unfortunately, we do not see that there is a way to quantify the effect on the CMI of any improvement in medical recordkeeping that may have occurred.

The rate of true change may have accelerated during the 1980s. Table 4.2 summarizes estimates of the rate of change available from the literature and the present study. Although the individual point estimate for 1987 to 1988 is relatively imprecise, its magnitude is quite consistent with the more precise 1986–1987 estimate. The combined studies suggest a rate of increase greater than the pre-PPS increase measured by Carter and Ginsburg (1985).

Table 4.2

Annual Rate of True Change in CMI

Period	Annual True Increase in CMI (%)	Source
1981–1983	0.5	Carter and Ginsburg (1985)
1983–1984	1.1	Carter and Ginsburg (1985)
1986–1987	1.6	Carter, Newhouse, and Relles (1990)
1987–1988	1.5	This study

NOTE: The standard errors of the 1986–1987 and 1987–1988 changes in the CMI were 0.50 percent and 0.77 percent of the CMI, respectively. Carter and Ginsburg (1985) did not report the standard error of their estimate of true change. Their estimate was calculated as a weighted average of several regression coefficients, and the standard errors of individual coefficients and the large sample size suggest that the statistical uncertainty associated with the 1981 through 1984 rates is substantially lower than that associated with later rates.

[1]Kahn et al. (1991) found an increase in the recording of laboratory tests on medical records between the 1981–1982 and 1985–1986 periods. We are aware of no analysis addressing whether medical records improved *during* the PPS period.

Our best estimate is that the 1987–1988 upward blip in the rate of increase in the CMI was caused entirely by coding and administrative factors, with the rate of true change remaining roughly constant. The increase seems to have been predominantly a result of coding improvements needed for accurate assignment of DRGs to cases under the new Grouper, rather than to DRG creep. These coding improvements can have salutary effects on nonfinancial uses of discharge abstract data, such as clinical research.

In FY 1990, the HCFA lowered the DRG weights by 1.22 percent to account for the "increase in the average case weight attributable to Grouper changes and recalibration between FY 1986 and FY 1988" (DHHS, 1989, p. 36472). This adjustment is modest in comparison to our estimate of the rate of increase in the CMI caused by coding and administrative factors. We estimate that such factors resulted in an increase of 1.7 percent from 1987 to 1988 (Table 4.1) and an increase of 0.8 percent from 1986 to 1987 (Carter, Newhouse, and Relles, 1990, Table 5.1), for a total increase of 2.5 percent. Thus, the HCFA's adjustment accounts for only one-half of the effect of coding and administrative factors during the 1986–1988 period on PPS expenditures in FY 1990 and later. Consistent with the PPS philosophy, no attempt was made to retroactively collect any of the extra expenditure in FY 1988 and FY 1989, which we estimate to be approximately $1 billion each year.

It seems clear to us that hospitals changed their coding of cases in response to the introduction of the 1988 Grouper. This response should be viewed as a predictable reaction to a change in the payment rules. We found that the SuperPRO also changed its coding behavior to more accurately assign the newly defined DRGs. These responses do, however, underscore the importance of considering likely coding changes before implementing future improvements in the Grouper, including the introduction of severity measures.[2]

In our view, the HCFA should routinely monitor the causes of CMI change. The size of the CMI increase in response to the 1988 Grouper emphasizes the importance of developing a method for ongoing monitoring of national increases in the CMI. The logical structure of our research shows the important elements of the analysis. All that is needed to implement CMI monitoring is a data base containing a nationally representative set of cases. The random sample cases selected by the PROs provide a way to easily construct such a data base.

---

[2]See, e.g., General Accounting Office (1988) or Health Systems Management Group (1988).

In our report on the 1986–1987 CMI increase (Carter, Newhouse, and Relles, 1990), we spelled out how the SuperPRO data base could be improved to serve this and other research purposes: (1) the SuperPRO could routinely identify the random sample cases within the PRO data bases; (2) a subsample of the random sample could be selected for each PRO, the subsample's size being proportional to the number of discharges in the area monitored by the PRO; and (3) the sample could include outlier cases and, to be even more useful, exempt hospitals and units as well.

Unfortunately, the SuperPRO has been given new PRO-monitoring instructions that may make it impossible to develop future stratified samples for monitoring the causes of CMI increases. It is our understanding that PRO operations will not be monitored annually in the future. Consequently, if the HCFA wishes to have an ongoing annual procedure for monitoring the sources of CMI increase, it will have to intervene to collect the necessary data now. In light of the money at stake, it seems to us that the HCFA should have this capability.

Appendix A

DECOMPOSITION OF CMI CHANGE USING 1987 GROUPER

The decomposition presented in the main body of the report uses the 1988 Grouper to measure how much of the CMI change is true change and how much is the effect of coding/administrative changes. If we use the FY 1987 Grouper to measure the effects of true change and coding/administrative changes, the Grouper normalization component must change so that the sum of all components still equals measured change. The Grouper normalization component appropriate for analysis with the 1987 Grouper is measured by applying both Groupers to the FY 1988 data rather than to the FY 1987 data. In the notation of Sec. 1 of this report, this Grouper normalization component is CMI88(H,C,G88) – CMI88(H,C,G87). For the other three components, one merely replaces G88 with G87 in the notation.

Table A.1 shows the results of the decomposition using the 1987 Grouper. This decomposition into true change versus change caused by coding and administrative factors is roughly similar to the equivalent decomposition using the 1988 Grouper (see Table 3.12). However, the contributions attributable to the different coding/

Table A.1

Components of CMI Change in Sample States Measured with 1987 Grouper

	Original Estimate		GLS Estimate	
Component	CMI Increase	Standard Error	CMI Increase	Standard Error
Measured change	0.03731	0.0012	0.03731	0.0012
True change	0.01863	0.0414	0.01230	0.0078
Coding/administrative changes				
Hospital only	0.00887	0.0071	0.00879	0.0071
Hospital and SuperPRO	0.00172	0.0033	0.00171	0.0033
Grouper normalization	0.01452	0.0004	0.01452	0.0004
Subtotal	(0.02511)	0.0077	(0.02502)	0.0077

administrative components of change are much different. The coding changes affected DRG assignment only when the 1988 Grouper was used. Thus, the effect of the coding components on the 1988 CMI change was much less when measured with the 1987 rather than the 1988 Grouper. The difference was counted under the Grouper normalization component instead of the coding component.

Appendix B

SUPPORTING TABLES

This appendix presents the details that support our decomposition of the CMI change.

Table B.1

Comparison of Hospital and PRO CMI Relative to SuperPRO Standard (Unweighted Analysis)

	Average DRG Weight Assigned by:			Increase over SuperPRO			
				Hospital		PRO	
Discharge Group	SuperPRO	Hospital	PRO	Average	Standard Error	Average	Standard Error
1987 Grouper							
1987, all states	1.19794	1.19735	1.19690	–0.00059	0.00363	–0.00104	0.00333
1987, sample states	1.20675	1.20810	1.20595	0.00135	0.00499	–0.00080	0.00464
1988, sample states	1.25870	1.26903	1.26503	0.01033	0.00622	0.00633	0.00522
1988 Grouper							
1987, all states	1.19460	1.18010	1.18141	–0.01451[a]	0.00417	–0.1319[a]	0.00386
1987, sample states	1.20372	1.19058	1.19289	–0.01314[a]	0.00561	–0.01083[a]	0.00516
1988, sample states	1.26413	1.26023	1.25810	–0.00390	0.00764	–0.00604	0.00667

[a] $p < .05$.

Table B.2

Regression of Weight Assigned by SuperPRO's Original Coding Minus Weight Assigned by PRO

PRO Coding Variables	FY 1987		FY 1988	
	Coefficient	t-Statistic	Coefficient	t-Statistic
Increase for complication or comorbidity	0.2444	10.64	0.2101	7.11
DRG 468	–0.2550	4.46	–0.1143	1.69
Surgical DRG	–0.0029	0.11	–0.0255	0.77
DRG weight	–0.0758	4.28	–0.0632	2.74
Weight for surgical cases	0.0390	2.00	0.0400	1.60
Intercept	0.0610	3.45	0.0543	2.34
R^2	0.11		0.06	

Table B.3

Regression of Weight Assigned by SuperPRO's New Coding Minus Weight Assigned by PRO (FY 1987 Cases)

PRO Coding Variables	Sample States		All States	
	Coefficient	t-Statistic	Coefficient	t-Statistic
Increase for complication or comorbidity	0.2090	6.79	0.2285	10.16
DRG 468	–0.2650	3.47	–0.1281	2.16
MDC 4, no ventilator	0.0343	1.61	0.0516	3.26
Surgical DRG	–0.0257	0.76	–0.0113	0.44
DRG weight	–0.0789	3.11	–0.0958	5.09
Weight for surgical cases	0.0158	0.57	0.0466	2.28
Intercept	0.0788	3.27	0.0948	5.32
R^2	0.07		0.06	

Table B.4

Estimated Effect of Inclusion of Outlier Cases on CMIs Calculated from SuperPRO Data

	1987 Grouper			1988 Grouper		
Sample CMI	Nonoutliers[a]	Outliers	Average	Nonoutliers[a]	Outliers	Average
Sample States						
Hospital/PRO, all cases						
FY 1987	1.20756	1.72542[b]	1.23190	1.19116	1.96547[b]	1.22754
FY 1988	1.26803	1.74383[b]	1.29295	1.25970	1.87395[b]	1.29186
SuperPRO, all cases, original coding						
FY 1987	1.20675	1.69606[c]	1.22974	1.20372	1.92314[c]	1.23752
FY 1988	1.25870	1.70201[c]	1.28191	1.26413	1.83999[c]	1.29428
SuperPRO, reabstraction sample						
Original coding	1.20767	1.69606[c]	1.23062	1.20392	1.92314[c]	1.23771
New coding	1.20966	1.69215[c]	1.23233	1.21307	1.92088[c]	1.24633
SuperPRO, real change sample[d]						
FY 1987	1.25043	1.69215[c]	1.27119	1.25560	1.92088[c]	1.28686
FY 1988	1.26704	1.70201[c]	1.28982	1.27283	1.83999[c]	1.30253
All States						
SuperPRO, reabstraction sample						
Original coding	1.19694	1.73888[c]	1.22240	1.19476	1.97213[c]	1.23129
New coding	1.20131	1.74707[c]	1.22695	1.20802	1.97892[c]	1.24424

[a]From SuperPRO sample.
[b]From MEDPAR.
[c]From regression.
[d]Population weighted.

REFERENCES

Carter, Grace M., and Paul B. Ginsburg, *The Medicare Case Mix Index Increase: Medical Practice Changes, Aging, and DRG Creep,* RAND, R-3292-HCFA, June 1985.

Carter, Grace M., Joseph P. Newhouse, and Daniel A. Relles, *How Much Change in the Case Mix Index is DRG Creep?* RAND, R-3826-HCFA, April 1990.

Department of Health and Human Services, "Medicare Program; Changes to the DRG Classification System," *Federal Register,* Vol. 52, No. 96, May 19, 1987, pp. 18877–18892.

Department of Health and Human Services, "Medicare Program; Changes to the Inpatient Hospital Prospective Payment System and Fiscal Year 1990 Rates," *Federal Register,* Vol. 54, No. 169, September 1, 1989, pp. 36452–36589.

General Accounting Office, "MEDICARE: Refinement of Diagnosis Related Groups Needed to Insure Payment Equity," GAO/HRD 88-41, April 1988.

Health Systems Management Group, *DRG Refinement with Diagnostic Specific Comorbidities and Complications: A Synthesis of Current Approaches to Patient Classification,* Final Report, Health Care Financing Administration Cooperative Agreements 15-C-98930/1-01 and 17-C-98930/1-02S1, School of Organization and Management, Yale University, New Haven, CT, 1988.

Kahn, Katherine L., et al., *The Effects of DRG-Based Prospective Payment System on Quality of Care for Hospitalized Medicare Patients,* RAND, R-3931-HCFA, forthcoming.

Prospective Payment Assessment Commission, "Medicare Prospective Payment and the American Health Care System: Report to the Congress," Washington, D.C., June 1989.

Simborg, D. W., "DRG Creep: A New Hospital-Acquired Disease," *New England Journal of Medicine,* Vol. 304, 1981, pp. 1602–1604.

Wonnacott, Ronald J., and Thomas H. Wonnacott, *Econometrics,* 2d ed., John Wiley & Sons, New York, 1979, pp. 427–428.

RAND/R-4098-HCFA/ProPAC